

1 **Manuscript Type:** Brief Report

2

3 **Title** Measuring the missing: greater racial and ethnic disparities in COVID-19 burden after  
4 accounting for missing race/ethnicity data

5

6 **Authors:** Katie Labgold\*<sup>1</sup> & Sarah Hamid\*,<sup>1</sup> Sarita Shah,<sup>1,2,3</sup> Neel R. Gandhi,<sup>1,2,3</sup> Allison  
7 Chamberlain,<sup>1</sup> Fazle Khan,<sup>4</sup> Shamimul Khan,<sup>4</sup> Sasha Smith,<sup>4</sup> Steve Williams,<sup>4</sup> Timothy L.  
8 Lash,<sup>1</sup> Lindsay J. Collin<sup>1,5</sup>

9

10 \*Co-First Authors

11

12 **Affiliations:** <sup>1</sup>Department of Epidemiology, Rollins School of Public Health, Emory University;  
13 <sup>2</sup>Department of Global Health, Rollins School of Public Health, Emory University; <sup>3</sup>Division of  
14 Infectious Diseases, Emory School of Medicine, Emory University, <sup>4</sup>Fulton County Board of  
15 Health; <sup>5</sup>Department of Population Health Sciences, Huntsman Cancer Institute, University of  
16 Utah;

17

18 **Corresponding Author:** Lindsay J Collin, Department of Population Health Sciences,  
19 Huntsman Cancer Institute, University of Utah, 2000 Circle of Hope Drive, Room 4746, Salt  
20 Lake City UT, 84112; email: [lindsay.collin@hci.utah.edu](mailto:lindsay.collin@hci.utah.edu)

21

22 **Running head:** Accounting for missing race/ethnicity COVID-19

23

24 **Conflicts of Interest:** The authors have no conflicts of interest to declare.

25

26 **Financial Support:** This work was supported in part by the US National Institutes of Health  
27 F31CA239566 (PI L. J. Collin), R01LM013049 (PI T. L. Lash), and K24AI114444 (PI N. R.  
28 Gandhi). It was also supported by a grant from the Robert W. Woodruff foundation (PI A.  
29 Chamberlain). K. Labgold is supported in part by the Center for Reproductive Health Research  
30 in the Southeast (RISE) Doctoral Fellowship and an ARCS Foundation Award.

31

32 **Data Access:** Due to patient confidentiality, data are only available upon request from the  
33 Fulton County Board of Health and with IRB approval from the Georgia Department of Public  
34 Health. Example code used to perform the imputation and bias-adjustment is available on  
35 GitHub.

36

37

38

39

40

41

42

43

44

45 **Abstract (147/150)**

46 Black, Hispanic, and Indigenous persons in the United States have an increased risk of SARS-  
47 CoV-2 infection and death from COVID-19, due to persistent social inequities. The magnitude of  
48 the disparity is unclear, however, because race/ethnicity information is often missing in  
49 surveillance data. In this study, we quantified the burden of SARS-CoV-2 infection,  
50 hospitalization, and case fatality rates in an urban county by racial/ethnic group using combined  
51 race/ethnicity imputation and quantitative bias-adjustment for misclassification. After bias-  
52 adjustment, the magnitude of the absolute racial/ethnic disparity, measured as the difference in  
53 infection rates between classified Black and Hispanic persons compared to classified White  
54 persons, increased 1.3-fold and 1.6-fold respectively. These results highlight that complete case  
55 analyses may underestimate absolute disparities in infection rates. Collecting race/ethnicity  
56 information at time of testing is optimal. However, when data are missing, combined imputation  
57 and bias-adjustment improves estimates of the racial/ethnic disparities in the COVID-19 burden.

58 **Keywords:** SARS-CoV-2, COVID-19, missing data, bias analysis, race/ethnicity disparities,  
59 surveillance  
60

## 61 Introduction

62 In the United States, early surveillance reports highlight that persons of Hispanic, Black, and  
63 American Indian/Alaskan Native race and ethnicity are disproportionately affected by the  
64 COVID-19 pandemic.<sup>1</sup> These disparities arise from historical and contemporary social and  
65 health inequities that result from systemic racism.<sup>2-4</sup> Racial capitalism in particular produces  
66 structurally unequal exposure to (and protection from) SARS-CoV-2 infection in key places of  
67 transmission (e.g. workplace).<sup>3</sup>

68 The role of systemic racism in the pandemic motivates the need for accurate surveillance of  
69 racial/ethnic disparities in SARS-CoV-2 infection and death. However, there are challenges in  
70 estimating COVID-19 racial/ethnic disparities.<sup>5,6</sup> Although reports highlight the unequal burden  
71 across racial/ethnic groups, the magnitude of disparities is uncertain because of the large  
72 proportion of missing race/ethnicity information in surveillance data. In recent reports,  
73 race/ethnicity information was missing in 56% of confirmed infections nationally and in 36% in  
74 Georgia.<sup>7,8</sup> Current surveillance estimates are reported as complete case analyses, which  
75 exclude cases with missing race/ethnicity.<sup>1,5,8,9</sup> Complete case analyses will bias racial/ethnic  
76 disparity estimates if race/ethnicity information is not missing completely at random.<sup>10</sup>

77 The Department of Health and Human Services issued COVID-19 reporting guidelines in June  
78 requiring all labs to report race/ethnicity beginning August 2020.<sup>11</sup> These guidelines seek to  
79 address missing data moving forward, but fail to address missing information for case-patients  
80 identified before August.

81 Collecting race/ethnicity information at time of testing is optimal, especially in surveillance of  
82 racial/ethnic health disparities. Until this becomes routine, imputation of missing race/ethnicity  
83 combined with quantitative bias-adjustment to account for misclassification of the imputed  
84 race/ethnicity can improve estimates of the COVID-19 burden among racial/ethnic groups when

85 race/ethnicity data are missing.<sup>12</sup> In this study, we calculate SARS-CoV-2 infection,  
86 hospitalizations, and case fatality rates by race/ethnicity group and report the absolute  
87 racial/ethnic disparities in SARS-CoV-2 infection rates in Fulton County, Georgia after  
88 accounting for missing race/ethnicity information.

89

## 90 **Methods**

91 Fulton County, Georgia, includes the city of Atlanta and residents identify as Black (44%), White  
92 (40%), Hispanic (7%), Asian (7%), and other races/ethnicities (2%).<sup>13</sup> Between 29 February  
93 2020 and 18 Aug 2020, 19,637 cases of SARS-CoV-2 infection were reported among Fulton  
94 County residents. Case reports included the patients' residential address, full name,  
95 race/ethnicity, hospitalization (yes/no/unknown), and death (yes/no/unknown). Fulton County  
96 Board of Health staff geocoded case-patients' address to census block groups. For this  
97 analysis, we categorized reported race/ethnicity as Black, Hispanic, Asian, White or Other.

98 We accounted for missing race/ethnicity information using a three-step approach: 1) imputation  
99 of race/ethnicity for all case-patients, 2) validation of the race/ethnicity imputation by calculating  
100 the accuracy of imputation among case-patients with reported race/ethnicity information, and 3)  
101 bias-adjustment of race/ethnicity estimates to account for misclassification of imputation among  
102 case-patients missing reported race/ethnicity information. Hereafter, we refer to race/ethnicity as  
103 reported when provided in case-patient records, *imputed* when referring to the imputed case-  
104 patient race/ethnicity, and *classified* when referring to the combined reported and imputed  
105 race/ethnicity after bias-adjustment.

106 First, for all case-patients we predicted their racial/ethnic group using the Bayesian Improved  
107 Surname Geocoding method.<sup>14</sup> This method estimates the probability of a person being  
108 classified as Black, Hispanic, Asian, White or Other race/ethnic group based on the case-

109 patient's surname and residential census block group, and the population distribution of  
110 race/ethnicity for census block groups and surnames. Imputation was performed using the R  
111 package "wru," which includes the 2010 surname census list with corresponding race/ethnicity  
112 distribution. Geographic distribution of race/ethnicity came from the 2018 5-year American  
113 Community Survey.<sup>15,16</sup> For the 546 (2.8%) case-patients who could not be geocoded,  
114 race/ethnicity was imputed using surname only.

115 Second, we validated the race/ethnicity imputation among case-patients whose race/ethnicity  
116 was available in the dataset (n=12,222, 64%). Predictive values (PV) were calculated for each  
117 imputed race/ethnic group. The PV is the probability that a person's reported race/ethnicity  
118 group classification was correctly imputed.<sup>12</sup>

119 Third, we used the PV values as bias parameters to quantitatively adjust for the expected  
120 misclassification of the imputed race/ethnicity groups. We assigned each race/ethnicity group  
121 PV from the validation to a Dirichlet distribution (**Table 1**). We then reclassified the imputed  
122 race/ethnicity probabilistically (100,000 iterations).<sup>12</sup> The quantitative bias-adjustment  
123 mathematically accounts for inaccurate assignment of case-patients to a race/ethnicity group by  
124 the Bayesian Improved Surname Geocoding method. Sampling error was incorporated into the  
125 estimates using bootstrap approximation from a standard normal distribution.<sup>12</sup>

126 For both the complete case and bias-adjusted analyses, we calculated the SARS-CoV-2  
127 infection rates (per 1,000 persons), hospitalization proportions (hospitalized cases/reported  
128 cases), and case fatality rates (deaths/reported cases) by race/ethnicity group. We reported  
129 95% confidence intervals (CI) for the complete case analysis and medians with 95% simulation  
130 intervals (SI) for the bias-adjusted estimates. We evaluated how accounting for missing  
131 race/ethnicity information impacts measures of racial/ethnic disparities by calculating the  
132 differences in SARS-CoV-2 infection rates in each race/ethnicity group compared with persons

133 of White race/ethnicity, among case-patients with reported race/ethnicity information, and  
134 among all case-patients after bias-adjustment. All analyses used R v3.6 (Vienna, Austria). The  
135 Georgia Department of Health determined this activity to be consistent with public health  
136 surveillance, so does not require informed consent or IRB approval.

## 137 **Results**

138 Among the 19,637 cases reported in Fulton County from 29 February to 19 August 2020, 7,145  
139 (36%) were missing race/ethnicity information in the case report. Data were more complete  
140 among the 1,840 hospitalized case-patients, where only 14 (3.5%) were missing race/ethnicity  
141 information. All deceased case-patients (n=456) had complete information on race/ethnicity.

142  
143 Comparison of reported versus imputed race/ethnicity group showed that the algorithm's  
144 imputation accuracy varied by imputed race/ethnicity group (**Table 1**). Of the 5,535 persons who  
145 were imputed as Black race/ethnicity, 93% (95%CI: 92%, 93%) were reported as Black in case  
146 reports (n=5,118). Among persons imputed as Hispanic ethnicity, 84% (95%CI: 82%, 85%)  
147 were reported as Hispanic. The algorithm was less accurate for case-patients with race/ethnicity  
148 imputed as Asian (PV=69%, 95%CI: 61%, 74%) and as White (PV=55%, 95%CI: 53%, 56%).  
149 The PV estimates for racial/ethnic groups changed over time, likely due to changes in the  
150 prevalence of demographic groups affected by the pandemic over time (**Supplemental Table**  
151 **1**).

152  
153 In both the complete case and bias-adjusted analyses, the SARS-CoV-2 infection rates were  
154 highest among those classified as Other, followed by Hispanic, Black, White, and Asian (**Table**  
155 **2a and 2b**). Imputation and bias-adjustment yielded higher estimates of infection rates than  
156 complete case analysis because more case-patients were included in the numerator. Estimated  
157 infection rates increased 1.8-fold for persons classified as Asian, 1.7-fold for White, 1.7-fold for

158 Hispanic, 1.6-fold for Other, and 1.5-fold for Black. Hospitalization proportions and case fatality  
159 rates decreased across all race/ethnicity groups with imputation and bias-adjustment compared  
160 with the complete case analyses, because more cases were included in the denominator. In  
161 both the complete case and bias-adjusted analyses, case-patients who were classified as Black  
162 race/ethnicity had the highest hospitalization proportions (complete case: 17%, 95%CI: 16%,  
163 18%; bias-adjusted: 12%, 95%SI: 11%, 12%) and case fatality rates (complete case: 4.6%,  
164 95%CI: 4.1%, 5.1%; bias-adjusted: 3.1%, 95%SI: 2.8%, 3.4%).

165  
166 The magnitude of the absolute disparity—difference in SARS-CoV-2 infection rates for case-  
167 patients classified in each race/ethnicity group compared with case-patients classified White—  
168 increased in the bias-adjusted analysis relative to the complete case analysis for nearly all  
169 race/ethnicity groups (**Table 3**). When comparing bias-adjusted with complete case results, the  
170 absolute disparity in infection rates increased 1.3-fold among classified Black and 1.6-fold  
171 among classified Hispanic race/ethnicity groups in reference to case-patients classified as  
172 White.

173

## 174 **Discussion**

175 In this study, accounting for missing race/ethnicity information revealed greater differences in  
176 SARS-CoV-2 infection rates comparing most racial/ethnic groups with case-patients of White  
177 race. These results suggest that national estimates, which exclude case-patients with missing  
178 race/ethnicity information, may underestimate the magnitude of absolute racial/ethnic disparities  
179 in COVID-19 morbidity and mortality.<sup>6,8</sup>

180 Our results underscore the need for imputation combined with bias-adjustment. In our study  
181 population, the PV estimates indicated that imputation alone overestimated infections among

182 case-patients classified as White and underestimated infections among case-patients classified  
183 as Black. Therefore, imputation alone would have been insufficient.

184 Both the complete case analysis and the bias-adjusted estimates demonstrate important  
185 absolute racial/ethnic disparities in the infection rates. The bias-adjusted estimates do not  
186 change our understanding of the direction of racial/ethnic disparities in the COVID-19 pandemic;  
187 however, the magnitude of racial/ethnic disparities changed meaningfully after bias-adjustment.  
188 In contrast, the hospitalization proportion and case fatality rate decreased across all classified  
189 race/ethnicity groups after accounting for missing race/ethnicity information because few  
190 hospitalized or deceased case-patients were missing race/ethnicity information. These results  
191 highlight the need for more complete reporting so that health equity and racial justice efforts  
192 aimed at addressing these disparities operate on the most accurate data possible.

193 The imputation of race/ethnicity has limitations. The Bayesian Improved Surname Geocoding  
194 algorithm limits the racial/ethnic groups that can be imputed to Black, Hispanic, Asian, White, or  
195 Other. The reliance on categories of 'other' is problematic for identifying and addressing  
196 disparities in other racial/ethnic populations (e.g. indigenous populations). Future studies should  
197 explore how accounting for missing race/ethnicity impacts other disease burden measures.

198 Our findings emphasize the importance of collecting complete race/ethnicity data at the time of  
199 testing, for the current pandemic and future outbreaks. When data are missing, Bayesian  
200 Improved Surname Geocoding combined with quantitative bias-adjustment provides better  
201 estimates of the racial/ethnic disparities in SARS-CoV-2 infection rates, hospitalization  
202 proportions, and case fatality rates.

203 **Tables**

**Table 1:** Predictive values (PV) and 95% confidence intervals (CI) of the imputation by race/ethnicity based on residence and surname compared with reported race/ethnic group in the State Electronic Notifiable Disease Surveillance System

		<b>Imputed Race/Ethnicity</b>				
		<b>Black</b>	<b>Hispanic</b>	<b>Asian</b>	<b>White</b>	<b>Other</b>
<b>Reported Race/Ethnicity</b>	<b>Black</b>	5118	68	13	1754	11
	<b>Hispanic</b>	77	1288	16	230	6
	<b>Asian</b>	16	15	145	80	4
	<b>White</b>	192	103	28	2827	2
	<b>Other</b>	132	68	12	302	1
	<b>Total</b>	5,535	1,543	214	5,193	24
	<b>PV % (95% CI)</b>	93% (92%, 93%)	84% (82%, 85%)	69% (61%, 74%)	55% (53%, 56%)	3.8% (0.1%, 15%)

204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227

228  
229

**Table 2a:** Complete case estimates of SARS-CoV-2 infection rates, hospitalization proportions, and case fatality rates by race/ethnic group among 12,222 cases reported to Fulton County Board of Health, 29 February – 18 Aug 2020.

Race/ Ethnicity	Total infections	Hospitalized	Died	At Risk*	Infection rate per 1,000 (95%CI)	Hospitalized percentage (95%CI)	Case Fatality Rate as a percentage (95%CI)
Asian	260	25	5	69987	3.7 (3.3, 4.2)	9.6 (6.2, 14)	1.9 (0.4, 3.8)
Hispanic	1617	214	15	74328	22 (21, 23)	13 (12, 15)	0.9 (0.5, 1.4)
Black	6964	1195	320	445992	16 (15, 16)	17 (16, 18)	4.6 (4.1, 5.1)
White	3152	312	112	406755	7.7 (7.4, 8.0)	9.9 (8.9, 11)	3.6 (2.9, 4.2)
Other	515	30	4	6056	85 (78, 92)	5.8 (3.9, 8.0)	0.8 (0.2, 1.6)

**Table 2b:** Bias-adjusted estimates of SARS-CoV-2 infection rates, hospitalization proportions, and case fatality rates including 7,415 cases with imputed race/ethnicity, among 19,637 cases reported to Fulton County Board of Health before 18 Aug 2020.

Race/ Ethnicity	Total infections (95%SI)	Hospitalized	Died	At Risk*	Infection rate per 1,000 (95%SI)	Hospitalized percentage (95%SI)	Case Fatality Rate as a percentage (95%SI)
Asian	456 (439, 474)	25	5	69987	6.5 (5.9, 7.2)	5.5 (3.4, 7.6)	1.1 (0.1, 2.1)
Hispanic	2,691 (2,661, 2721)	214	15	74328	36 (35, 38)	7.9 (6.9, 9.0)	0.6 (0.3, 0.8)
Black	10,838 (10,327, 10,428)	1195	320	445992	23 (23, 24)	12 (11, 12)	3.1 (2.8, 3.4)
White	5,303 (5,250, 5,356)	312	112	406755	13 (13, 13)	2.1 (1.7, 2.5)	2.1 (1.7, 2.5)
Other	837 (810, 865)	30	4	6056	138 (128, 148)	0.5 (0.0, 0.9)	0.5 (0.0, 0.9)

\*American Community Survey 5-year 2018 estimates

230  
231  
232  
233  
234  
235  
236  
237

**Table 3:** Relative difference (RD) of SARS-CoV-2 infection rates among minority groups compared with non-Hispanic White persons among cases with complete information and after accounting for missing race/ethnicity among 4004 SARS-CoV-2 infected persons reported to Fulton County before 20 May 2020.

Race/ Ethnicity	Complete Case		Bias-Adjusted		Relative change in magnitude of disparity
	Infection rate per 1,000 (95%CI)*	RD per 1,000 (95%CI)	Infection rate per 1,000 (95%SI)	RD per 1,000 (95%SI)	
Asian	3.7 (3.3, 4.2)	-4.0 (-4.6, -3.5)	6.5 (5.9, 7.2)	-6.5 (-6.8, -6.2)	0.6
Hispanic	22 (21, 23)	14 (13, 15)	36 (35, 38)	23 (23, 23)	1.6
Black	16 (15, 16)	7.9 (7.4, 8.3)	23 (23, 24)	10 (10, 11)	1.3
White	7.7 (7.4, 7.8)	Reference	13 (13, 13)	Reference	
Other	85 (78, 92)	77 (70, 84)	138 (128, 148)	125 (121, 130)	1.6

238

239

240 **Appendix**

241

**Supplemental Table 1:** Positive predictive value (PPV) of the imputation by race/ethnicity based on residence and surname compared with the reported race/ethnic group in COVID-19 case report stratified by months (March through May and June through August) of diagnosis

		Predicted Race/Ethnicity				
		Asian	Black	Hispanic	Other	White
<b>March–May</b>						
Reported Race/Ethnicity	<b>Asian</b>	33	2	6	2	28
	<b>Black</b>	3	1183	12	3	654
	<b>Hispanic</b>	1	7	156	0	38
	<b>Other</b>	9	17	11	0	45
	<b>White</b>	5	69	30	0	608
	<b>PPV</b>	65%	93%	73%	0%	44%
<b>June–Aug</b>						
Reported Race/Ethnicity	<b>Asian</b>	112	14	9	2	52
	<b>Black</b>	10	3935	56	8	1100
	<b>Hispanic</b>	15	70	1132	6	192
	<b>Other</b>	3	115	57	1	257
	<b>White</b>	23	123	73	2	2219
	<b>PPV</b>	69%	92%	85%	5%	58%

242

243

244

245

246 **References**

- 247 1. Stokes EK, Zambrano LD, Anderson KN, et al. Coronavirus Disease 2019 Case  
248 Surveillance - United States, January 22-May 30, 2020. *MMWR Morb Mortal Wkly Rep.*  
249 2020;69(24):759-765. doi:10.15585/mmwr.mm6924e2
- 250 2. Health Equity Considerations & Racial & Ethnic Minority Groups. National Center for  
251 Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases.  
252 [https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-](https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html)  
253 [minorities.html](https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html). Published 2020. Accessed July 17, 2020.
- 254 3. McClure ES, Vasudevan P, Bailey Z, Patel S, Robinson WR. Racial Capitalism within  
255 Public Health: How Occupational Settings Drive COVID- 19 Disparities. *Am J Epidemiol.*  
256 2020:113-120.
- 257 4. Egede LE, Walker RJ. Structural Racism, Social Risk Factors, and Covid-19 — A  
258 Dangerous Convergence for Black Americans. *N Engl J Med.* 2020;383(12):e77(1)-  
259 e77(3). doi:10.1056/NEJMp2023616
- 260 5. Servik K. 'Huge hole' in COVID-19 testing data makes it harder to study racial disparities.  
261 *Science (80- )*. July 2020. doi:0.1126/science.abd7715
- 262 6. Cowger TL, Davis BA, Etkins OS, et al. Comparison of Weighted and Unweighted  
263 Population Data to Assess Inequities in Coronavirus Disease 2019 Deaths by  
264 Race/Ethnicity Reported by the US Centers for Disease Control and Prevention. *JAMA*  
265 *Netw open.* 2020;3(7):e2016933. doi:10.1001/jamanetworkopen.2020.16933
- 266 7. Georgia Department of Public Health COVID-19 Daily Status Report.  
267 <https://dph.georgia.gov/covid-19-daily-status-report>. Published 2020. Accessed July 18,  
268 2020.
- 269 8. Oppel R, Gebelhoff R, Lai K, Wright W, Smith M. The Fullest Look Yet at the Racial  
270 Inequity of Coronavirus. *New York Times*.  
271 <https://www.nytimes.com/interactive/2020/07/05/us/coronavirus-latinos-african->

- 272 americans-cdc-
- 273 [data.html?campaign\\_id=2&emc=edit\\_th\\_20200706&instance\\_id=20039&n](https://www.cdc.gov/data.html?campaign_id=2&emc=edit_th_20200706&instance_id=20039&n)
- 274 [l=todaysheadlines&regi\\_id=71026656&segment\\_id=32674&user\\_id=c99fb](https://www.cdc.gov/data.html?campaign_id=2&emc=edit_th_20200706&instance_id=20039&n)
- 275 [3a6b3b754c](https://www.cdc.gov/data.html?campaign_id=2&emc=edit_th_20200706&instance_id=20039&n). Published 2020. Accessed July 18, 2020.
- 276 9. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2
- 277 infection in the United States. *Nat Commun*. 2020. doi:10.1038/s41467-020-18272-4
- 278 10. Perkins NJ, Cole SR, Harel O, et al. Principled Approaches to Missing Data in
- 279 Epidemiologic Studies. *Am J Epidemiol*. 2017;187(3):568-575. doi:10.1093/aje/kwx348
- 280 11. *The Coronavirus Aid, Relief, and Economic Security (CARES) Act*. United States; 2020.
- 281 12. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*.
- 282 (Gail M, Krickeberg K, Samet J, Tsiatis A, Wong W, eds.). New York: Springer; 2009.
- 283 doi:10.1007/978-0-387-87959-8
- 284 13. American Community Survey: Hispanic or Latino Origin by Race. The United States
- 285 Census Bureau. <https://data.census.gov/cedsci/table?t=Race and>
- 286 [Ethnicity&g=0500000US13121&tid=ACSDT5Y2018.B03002&moe=false&hidePreview=true](https://data.census.gov/cedsci/table?t=Race and)
- 287 [ue](https://data.census.gov/cedsci/table?t=Race and). Published 2020. Accessed August 19, 2020.
- 288 14. Elliott MN, Fremont A, Morrison PA, Pantoja P, Lurie N. A new method for estimating
- 289 race/ethnicity and associated disparities where administrative records lack self-reported
- 290 race/ethnicity. *Health Serv Res*. 2008;43(5 P1):1722-1736. doi:10.1111/j.1475-
- 291 [6773.2008.00854.x](https://doi.org/10.1111/j.1475-6773.2008.00854.x)
- 292 15. About the American Community Survey. US Census Bureau.
- 293 <https://www.census.gov/programs-surveys/acs/about.html>. Published 2020. Accessed
- 294 July 8, 2020.
- 295 16. Khanna K, Imai K. Package ‘wru’: Who are You? Bayesian Prediction of Racial
- 296 Category Using Surname and Geolocation. 2019. [https://cran.r-](https://cran.r-project.org/web/packages/wru/wru.pdf)
- 297 [project.org/web/packages/wru/wru.pdf](https://cran.r-project.org/web/packages/wru/wru.pdf).

